Experimental Legal Studies on Perceived Procedural Justice and Trust in Law and Society

Hulst, J.E.

2017

document version
Publisher's PDF, also known as Version of record

4

Trust in Law and Society: An Experimental Approach to Interviewer Effects Among Under-Investigated Participants

This chapter is based on an ongoing line of research included in: Hulst, L., Van den Bos, K., Robijn, M., Romijn, S., Schroen, S., & Wever, Th. (2017). *Trust in law and society: An experimental approach to interviewer effects among under-investigated participants.* Manuscript submitted for publication.

Author contributions: Liesbeth Hulst co-designed Studies 4.1 and 4.2, commented on the design of Study 4.3, conducted statistical analyses, interpreted results, and wrote the manuscript. Kees van den Bos co-designed the experiments, provided conceptualization and theory used to integrate findings, conducted statistical analyses and interpreted results, and edited the manuscript. Marianne Robijn aided in designing Study 4.1 and collected the data of Study 4.1 for her Msc. thesis. Sietske Romijn aided in designing Study 4.2 and collected the data of Study 4.2 for her Msc. thesis. Sharon Schroen aided in designing Study 4.3 and collected the data of Study 4.3 for her Msc. thesis.

Abstract

Extending observations indicating that the behavioral sciences tend to rely heavily on Western, Educated, Industrialized, Rich, and Democratic (WEIRD) participants, the current paper argues that studies on trust in law and society may be missing crucial patterns of participant reactions because participants are tested by WEIRD interviewers. Field experiments designed specifically to test this assumption show that when answering questionnaires on degree of trust in law and society as given to them by interviewers presenting themselves as coming from Law Schools, lower educated people indicated that they hold high levels of trust in Dutch judges. That pattern replicates a finding that is often seen in trust surveys. Yet, when the same interviewer presented herself as coming from a lower educated background, participants reported much less trust in judges. Similar effects were found among female Muslims who indicated higher trust in Dutch society when interviewed by an interviewer who was not wearing a headscarf as opposed to was wearing a headscarf. Taken together, these findings point at the importance of experimentally investigating the pivotal role that interviewers can play in studies on trust in law and society, especially when examining under-investigated participants such as those with lower education or female Muslims.

Trust in Law and Society: An Experimental Approach to Interviewer Effects Among Under-Investigated Participants

People's trust in law and society is an important indicator of the state of democracy (De Tocqueville, 1945). It is also indicative of the legitimacy of a given society and important institutions pertaining to law and society (e.g., Tyler & Huo, 2002). For example, it is assumed that it is through trust that citizens confer important societal institutions such as the judiciary with the legitimacy that the justice system needs to operate effectively (Tyler, 2006; Tyler & Jackson, 2014). Hence, public trust in law and society is considered important for social cohesion and for the effective functioning of societal institutions and the broader society (e.g., Dohmen, Verbakel, & Kraaykamp, 2010; Hough, Jackson & Bradford, 2013; Tam, Hewstone, Kenworthy, & Cairns, 2009; Te Riele & Roest, 2009). It should not come as a surprise, therefore, that the issue of trust in law and society serves a prominent role in many surveys and other research projects (e.g., TNS Political & Social, 2013) and is a key issue in the study of law and human behavior.

In the current paper we study trust in Dutch judges among lower educated, intermediate educated, and higher educated participants (Studies 4.1 and 4.2). We also assess trust in Dutch society (including trust in majority groups) among female Muslim participants (Study 4.3). We predict that the interviewer who conducts the trust study can have important effects on the degree of trust in law and society our participants report. We test this prediction in three experimental studies that we conducted in field settings. In these studies we experimentally vary potentially interesting differences between interviewers and examine how these manipulated interviewer differences affect participants' trust in law and society. One source of inspiration for the current line of research comes from the concept of "WEIRD" participants (Henrich, Heine, & Norenzayan, 2010), which we want to extend to the potentially relevant issue of "WEIRD" interviewers.

WEIRD interviewers

Henrich et al. (2010) observed that the vast majority of studies in the behavioral sciences rely heavily on participants from Western, Educated, Industrialized, Rich, Democratic, in short WEIRD, backgrounds. In fact, in the psychology studies reviewed by Henrich et al. (2010) 96% of participants were from Western industrialized countries. And even within the West, 67 to 80% of the samples of experimental studies were composed solely of undergraduates in psychology courses (Arnett, 2008; Henrich et al., 2010). Other behavioral science studies in the field of cognitive science, economics, and decision science are also dominated by undergraduate samples (Henrich et al., 2010).

Populations vary considerably on important aspects of psychology, motivation, and behavior (Henrich et al., 2010). Interestingly, Henrich at al. (2010) note that there is substantial variability in experimental results across populations and that WEIRD participants are particularly unusual when compared with the rest of the species. So, the fact that the vast majority of studies use WEIRD participants presents a challenge to the understanding of human psychology and behavior (Henrich et al., 2010). Here we argue that this has important implications for related disciplines such as the field of law and human behavior that have been incorporating insights from psychology and social and behavioral sciences and that have been examining how these insights can translate into policy and decision making.

In this paper we propose that WEIRD-ness does not just apply to WEIRD participants, but is also relevant to the concept of "WEIRD interviewers" that we introduce here. That is, we observe that research studies into trust in law and society are often conducted by interviewers from Western, Educated, Industrialized, Rich, Democratic backgrounds. For example, many interview studies reported in *Law and Human Behavior* are conducted by interviewers coming from Law Schools or other university departments. We note that in some studies interviewers definitely come from non-WEIRD backgrounds. Quite often though, interviewers who examine trust in law and society will be highly educated and come from majority groups. Thus, interviewers will often be WEIRD. Furthermore, if interviewers are from non-WEIRD backgrounds this aspect is not systematically included in the research design of the studies involved. Following these observations, we wondered whether the WEIRD-ness of interviewers might have an effect on participants' report of trust in law and society.

Specifically, in the current paper we wanted to examine how the backgrounds of interviewers affect trust reports of groups of under-investigated participants. In particular, we focus on trust in law and society as perceived by lower educated participants and female Muslim participants. These are participants who have remained relatively under-investigated in many research studies (Bovens, Dekker, & Tiemeijer, 2014; Huijnk, Dagevos, Gijsberts, & Andriessen, 2015). We argue that understanding trust in law and society among these participants is an important issue and that studying how these participants respond to interviewers coming from university or non-university departments or interviewers from majority (non-Muslim) or minority (Muslim) groups will reveal important insights into the issue of trust in law and society.

Social Distance

Besides the reasons mentioned earlier, the studies that we are presenting here are important, we think, because they can be related to social distance between different groups in society, such as those with lower versus higher educational levels

and between members of majority versus non-majority groups. That is, the current studies are also inspired by a development of societal polarization that can be observed in many different societies. For example, in the field of political science it has been argued that polarization has developed between societal groups of people with different social-cultural orientations, and these polarization processes are linked with people's educational level (Bovens et al., 2014; Kriesi et al., 2008; Kriesi et al., 2012; Stubager 2010) and with people's ethnic or religious backgrounds (e.g., Moors, Balogh, Van Donselaar, & De Graaff, 2009).

Specifically, with respect to education-related polarization, it has been argued that there are groups consisting of are mostly higher educated people who are reasonably positive about globalization issues (such as open borders, other cultures and admitting immigrants), who tend to have high political trust, and who are internationally oriented and frequently live in cities. This can be contrasted with groups consisting of often lower educated people who tend to focus on the disadvantages of open borders and immigration, have low political trust, are more locally oriented, and more often live outside cities (see Bovens et al., 2014; Kriesi et al., 2008; Kriesi et al., 2012; Stubager, 2010).

It has further been argued that high and low educated people increasingly tend to live in separate social environments, with low educated persons having few interactions with high educated persons (Bovens et al., 2014; Stubager, 2010). Thus, polarization and segregation has developed between lower and higher educated people (see Bovens et al., 2014).

Within many Western societies there are also processes of polarization between Muslims and the majority group of non-Muslims (see, e.g., Moors et al., 2009; Tajfel, & Turner, 1979). This may express itself in negative attitudes about each other (Van den Bos, Loseman, & Doosje, 2009; Vedde, Van Geel, Wenink, & Zwaanswijk, 2012), increasing discrimination against Muslims (Andriessen, Fernee, & Wittebrood, 2014), and the experience of status differences between Muslims and non-Muslims (Vrooman, Gijsberts, & Boelhouwer, 2014). Thus, these polarization processes are linked with ethnic or religious backgrounds (Moors et al., 2009).

These developments of societal polarization and segregation suggest that there is social distance between members of these groups in society. That is, there may be social distance between high and low educated people and between Muslims and non-Muslims in at least some Western countries. Social distance basically means how close we feel to another person (Aron, Aron, & Smollan, 1992), and the concept has been used successfully in investigating the strength of intimate and other social relationships (e.g., Aron et al., 1992). The experience of social distance has also been shown to be sensitive to intergroup relations such that people feel closer to their ingroup members

than to outgroup members (Schubert & Otten, 2002). Interestingly, people perceive whether someone is close or distant from them in spontaneous manners, even when social distance is not directly related to their current goals (Bar-Anan, Liberman, Trope, & Algom, 2007). In fact, social distance is considered to be a fundamental processing principle that guides people's cognition and action (Bar-Anan et al., 2007).

We propose that the issue of social distance may be relevant for how studies on trust in law and society are commonly conducted. We suggest that the way in which interviewers on trust in law and society often present themselves may for some participants activate the impression that the interviewer belongs to a different social category or group than these participants are used to and themselves belong to. Hence, the way in which interviewers present themselves may for some participants prompt that they experience social distance between the interviewer and themself.

Interviewers on trust in law and society will often come from organizations that are part of the societal status quo such as Law Schools or other university departments, or at least they might be perceived as part of the status quo by some kinds of participants, such as lower educated people. We note that sometimes interviewers do come from different types of organizations that are not part of the status quo, and we note that differences between interviewers' affiliation tend not be a systematic part of the research designs of trust studies. We wonder whether such potentially interesting differences between interviewers affect reported levels of trust in law, especially among lower educated participants.

Similarly, often, although not always, interviewers on trust in law and society will belong to the majority group of the society in which the research study takes place. In Western countries, this means that interviewers will often *not* be Muslim. We wonder whether female Muslim participants will report different levels of trust in law depending on whether the female interviewer is presenting herself with or without a headscarf.

The Current Research

In this paper we systematically examine how potentially interesting differences between interviewers affect trust in law and society among under-investigated participants. We present three controlled experimental studies that tested this idea. We used a comparable set-up in all studies. All three studies included an experimental manipulation that varied how the interviewer presented herself to the participants.

In the first two studies, the interviewer introduced herself to the participants as coming from an university Law School in some conditions of our experiment. We contrasted this with conditions in which the same interviewer presented herself as coming from a non-university organization. In Study 4.2 we also included a condition in which the interviewer presented herself as coming from an university Psychology Department.

Interviewers on trust in law and society are often coming from a Law School or similar type of organization. Hence, we suggest that the Law School interviewer conditions represent how interviewers frequently present themselves in studies on trust in law and society. And we note that Law Schools (and to a lesser extent Psychology Departments) tend to be part of the societal status quo. We also note that if interviewers are coming from non-university type of organizations, the background of the interviewers tends not be a systematic part of the design of studies on trust in law and society and tends not be varied by means of experimental manipulation.

In our first two studies we crossed our interviewer manipulations with educational levels as defined by the official Dutch Central Bureau of Statistics (CBS, 2008). The CBS distinguishes three categories of educational levels that in the Netherlands are qualitatively different from each other. Thus, following this official classification, we distinguished between participants with lower, intermediate, and higher educational levels in Studies 4.1 and 4.2 and examined what happens when the university Law School interviewer versus the non-university interviewer (versus the university Psychology interviewer in Study 4.2) asks lower educated, middle educated and higher educated participants to respond to a scale assessing their trust in judges at the institutional level. On the basis of what we have proposed here, we were particularly interested in whether lower educated participants would indicate different levels of trust in judges toward the same person presenting herself as a Law School interviewer or as a non-university interviewer.

Study 4.3 included a condition in which the interviewer presented herself without a headscarf to female Muslim participants. We contrasted this with a condition in which the same interviewer presented herself with a headscarf to our female Muslim participants. Interviewers conducting studies on trust in law and society are mostly coming from majority groups. This implies that in Western countries trust survey studies are mostly conducted by interviewers who do not wear a headscarf. Thus, we suggest that the no-headscarf interviewer condition represents how interviewers commonly present themselves in studies on trust in law and society. Furthermore, we note that if interviewers are from minority groups the background of the interviewers generally is not included in the design of the studies and is not systematically varied by means of experimental design.

In Study 4.3, we compared female Muslim participants' responses on issues of trust in society and trust in non-Muslims when the interviewer presented herself without a headscarf versus with a headscarf to female Muslim participants. The question that we assessed in Study 4.3 is whether female Muslims indicate different levels of trust in Dutch society and trust in non-Muslims toward an interviewer without a headscarf versus with a headscarf.

We conducted all three experiments in field settings. That is, participants were approached at indoor shopping malls and train stations in the Netherlands. These are places that are visited by people from different social demographic backgrounds and where we obtained heterogeneous samples of participants. In all studies, the interviewer used standard, pre-scripted communication in her interaction with participants. Sampling procedures were the same across interviewer conditions. We interviewed the participants thus obtained by means of structured, paper-and-pencil questionnaires that participants completed in anonymous ways. Furthermore, various measures were taken to minimize possible experimenter demands in our studies (Orne, 1962).

Study 4.1

In Study 4.1 we compared participants' responses on a scale of trust in Dutch judges toward an interviewer that presented herself as coming from Utrecht University Law School with responses when the same interviewer presented herself as coming from (the non-university) Utrecht Regional Community College. With this experimental method we wanted to examine whether participants respond differently to a Law School interviewer than to a Regional Community College interviewer when asked about their trust in the judges in their country.

In Study 4.1, we crossed the interviewer manipulation with three categories of educational levels that are qualitatively different, at least in the Netherlands, as defined by the official Dutch Central Bureau of Statistics (CBS, 2008). We were particularly interested in whether lower educated participants would indicate different levels of trust in judges to the Law School interviewer than to the Regional Community College interviewer.

Method

Participants and design. One hundred and twenty-six participants took part in Study 4.1. Participants were interviewed by the same interviewer presenting herself as coming from Utrecht University Law School (N = 62) or Utrecht Regional Community College (N = 64). There were 40 participants with low educational levels, 32 with intermediate educational levels, and 54 with high educational levels, in accordance with the division in three qualitatively different educational levels in the Netherlands (CBS, 2008). A total of 52 men and 74 women participated voluntarily in the study. Participants had an average age of 46.3 years (SD = 14.8; range: 18-83 years). Data were collected on week days between February 18 and March 20, 2014.

The key prediction in this study was an interaction effect between the interviewer manipulation and educational level. A post-hoc G-power analysis (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that with $\alpha = .05$ and a medium-large effect size of the predicted interaction effect ($f = .325$) the sample of Study 4.1 has a statistical power

of .90 to detect the predicted interactive effect of interviewer and educational level. A sufficient amount of power is at least .80 (Cohen, 1992).

Experimental procedure.

Interviewer manipulation. In the *Law School condition* of Study 4.1 the interviewer introduced herself as Marie-Claire from Utrecht University Law School. As can be seen in the left part of Figure 4.1, the interviewer in this condition dressed accordingly as a law student who is conducting interviews. She wore neat clothing, including a jacket, skirt, and shoes with small heels.

In the *non-university condition*, the same interviewer introduced herself as Priscilla from a non-university Regional Community College. As can be seen in the right part of Figure 4.1, in this condition the interviewer dressed informally, wearing casual clothing like jeans, a track jacket, and sneakers.

Figure 4.1. The interviewer manipulation of Study 4.1 illustrated: The interviewer introduced as coming from Utrecht University Law School (left) or Utrecht Regional Community College (right).

Procedure. Our experimenter went out to a shopping mall in Utrecht and asked people from various backgrounds and with different educational levels to complete a questionnaire about their trust in judges. People from different social demographic backgrounds visited the mall.

One interviewer condition was used per day and conditions alternated each day. To reduce selection bias, sampling procedures were kept similar between interviewer conditions and the type of days and hours at which the interviewer visited the mall were the same across interviewer conditions. Importantly, the experimenter asked every third person who passed by to participate in the study, without distinguishing on passers-by's apparent educational level, age, or gender. The experimenter used standard pre-scripted communication in the interaction with participants. Response rates did not differ between interviewer conditions, which to us suggested that we were successful in our aim to minimize selection effects.

The experimenter interviewed the participants thus obtained by means of structured, paper-and-pencil questionnaires that participants could complete in anonymous ways. The experimenter made sure that participants were free to fill the questionnaires on their own, without the experimenter being in their immediate presence. The experimenter took several meters distance after handing out the questionnaires and remained at this distance while participants completed their questionnaires.

Furthermore, the experimenter gave participants envelopes with the questionnaires inside, and made sure that after filling out the questionnaires, participants put the questionnaires back into the envelopes, sealed the envelopes, and only then gave the materials back to the experimenter. The experimenter used standard pre-scripted communication in the interaction with participants to reduce the potential influence of experimenter demands.[1] Finally, the experimenter conducted debriefing interviews, which indicated that we were successful in our aim to rule out potential experimenter demands (Orne, 1962).

Trust in judges. Our main dependent variable was a measure of trust in the judges in the Netherlands, which we based on our analysis of literature on trust (e.g., Rotter, 1967; Rousseau, Sitkin, Burt, & Camerer, 1998; Schoorman, Mayer, & Davis, 1996), including trust in legal authorities and the specific context of public trust we are studying here (e.g., Tyler & Huo, 2002; Tyler & Jackson, 2014). Specifically, we assessed trust in judges in the Netherlands by asking participants to what extent they agreed with the following statements: "I have trust in the Dutch legal system," "I have the feeling that the judges in the Netherlands can be trusted," "I am happy with the judges in the Netherlands," "Dutch judges protect the interests of common people," and "Dutch judges understand

[1] The exact scripts that our interviewers used are available on request.

the problems of common people."[2] Ratings were made on 7-point Likert-type scales (1 = not at all, 7 = very much). Participants' answers to these items were internally consistent ($\alpha = .92$) and average scores were used for interpretation, with higher scores indicating higher levels of trust in judges.

Results

Trust in judges. To test our predictions, we performed a 2 (interviewer: Law School, Regional Community College) × 3 (educational level: low, intermediate, high) analysis of variance (ANOVA) on trust in judges. This analysis revealed a significant main effect of the interviewer manipulation, $F(1, 120) = 7.71$, $p < .01$, $\eta_p^2 = .06$, a significant main effect of educational level, $F(2, 120) = 10.13$, $p < .001$, $\eta_p^2 = .14$, and a significant interaction effect between the interviewer manipulation and participants' educational levels, $F(1, 120) = 6.97$, $p < .01$, $\eta_p^2 = .10$. To interpret these effects, we performed a post-hoc Tukey-b test for multiple comparisons between the means ($p = .05$) with the six cells of our design serving as the independent variable (Kirk, 1982). Table 4.1 shows the results of this test, and the means and standard deviations of participants' trust in judges.

As can be seen in Table 4.1, the Tukey-b test showed that when participants were interviewed by the Law School interviewer trust in judges among lower educated participants was as high as among participants with intermediate or high educational levels. However, when the same interviewer presented herself as coming from the Regional Community College, lower educated participants' trust in judges was significantly lower compared to the trust levels lower educated participants indicated to the Law School interviewer, and compared to trust levels among intermediate and high educated participants. Participants with intermediate and high education did not indicate different trust levels as a function of whether the interviewer presented herself as from the Law School or the Regional Community College. Thus, while trust in judges indicated to the Law School interviewer was relatively high for participants with all

[2] We report all manipulations, all data exclusions, and all measures in our studies, so we note that we focused the current studies on trust. In the first study, after measuring trust first, we also assessed different types of reactions. The items not reported were included for exploratory purposes, did not affect the effects reported, and are available on request. Although outside the scope of our current hypothesis, we also included as an exploratory element in Study 4.1 (and not Studies 4.2 and 4.3) an additional factor in the design in which we asked or did not ask participants to recall past disinhibited behavior (in line with earlier work; see Van den Bos, Müller, & Van Bussel, 2009; Van den Bos, Van Lange et al., 2011). There was no significant main effect or significant interaction effect of this manipulation on participants' trust judgments. This factor was therefore dropped from the analyses reported here so that in all studies reported in this paper we focus on comparing trust in law and society between interviewer conditions. Complete details and results are available on request.

Table 4.1.
Trust in judges as a function of the interviewer being introduced as coming from Utrecht University Law School or Utrecht Regional Community College and participants' educational level (Study 4.1)

	Educational Level					
	Low		Intermediate		High	
Interviewer	*M*	*SD*	*M*	*SD*	*M*	*SD*
Law School	$4.69_{a,b}$	1.33	3.80_{a}	1.06	5.03_{b}	1.34
Regional Community College	2.72_{c}	1.50	$4.03_{a,b}$	1.34	$4.72_{a,b}$	1.27

Note. Trust ratings are on 7-point scales with higher values indicating higher levels of trust in Dutch judges. Means with no subscript in common differ significantly ($p < .05$), as indicated by a Tukey-b post-hoc test for multiple comparisons between means among the six cells of our design (Kirk, 1982).

educational levels, lower educated participants' trust in judges was significantly lower when interviewed by the Community College interviewer.

Discussion

In summary, most of the action of our interviewer manipulation took place among lower educated participants. That is, the trust levels that lower educated participants indicated to the university (Law School) interviewer were as high as trust levels among middle and high educated participants. However, when lower educated participants were interviewed by the non-university (Regional Community College) interviewer, lower educated participants indicated significantly lower levels of trust in judges.

Study 4.2

To get an indication of the robustness of the effects studied here, we had a different experimenter who held the interviews of Study 4.2. This study was conducted at train stations and we used different university and non-university affiliations. In Study 4.2 we also added an extra interviewer condition to explore how participants would react to an interviewer who introduced herself as coming from a university Psychology Department. Thus, in Study 4.2, we had three interviewer conditions. First, we had a condition in which the interviewer presented herself as coming from Leiden University Law School. Second, we had the newly added condition in which the interviewer presented herself as coming from Leiden University Psychology Department. Third, in the non-university condition, the same interviewer presented herself as coming from Leiden Regional

Community College. The Law School and Regional Community College conditions were similar to Study 4.1.

The main aim of Study 4.2 was to examine whether especially lower educated participants would again respond differently in terms of trust in judges toward an interviewer that presented herself as coming from a Law School compared to when the same interviewer presented herself as coming from a Regional Community College. In Study 4.2, we also examined how participants would respond to the interviewer when she presented herself as coming from a Psychology Department.

Method

Participants and design. One hundred and forty-three participants took part in Study 4.2. Participants were interviewed by the same interviewer presenting herself as coming from Leiden University Law School ($N = 48$), Leiden University Psychology Department ($N = 47$), or Leiden Regional Community College ($N = 48$). In the sample, there were 41 individuals with lower educational level, 45 with intermediate educational levels, and 53 with high educational levels (CBS, 2008; 4 scores were missing). A total of 68 men and 75 women participated voluntarily in the study. Participants had an average age of 37.36 years ($SD = 15.19$; range: 16-74 years). Data were collected on weekdays between December 3, 2014 and January 14, 2015.

The key prediction in this study was an interaction effect between the interviewer manipulation and educational level. The sample size was based on an a-priori G-power analysis (Faul et al., 2007) with $\eta_p^2 = .10$ based on the interaction effect from Study 4.1, $\alpha = 0.05$, and high statistical power of .90.

Experimental Procedure.

Interviewer manipulation. In Study 4.2, we again had an interviewer manipulation, this time with three conditions. The Law School and Community College conditions were similar to Study 4.1. Thus, in the *Law School condition*, the interviewer introduced herself as Marie-Claire from Leiden University Law School. As can be seen in the left part of Figure 4.2, the interviewer dressed accordingly as a law student who is conducting interviews, wearing neat clothing like a jacket, skirt, and shoes with small heels.

Also similar to Study 4.1, in the *non-university condition*, the same interviewer presented herself as coming from Leiden Regional Community College. As can be seen in the right part of Figure 4.1, she dressed informally wearing casual clothing like jeans, a track jacket, and sneakers. The interviewer visited the same train stations in all interviewer conditions.

In the *Psychology Department condition* the interviewer introduced herself as Inge from Leiden University Psychology Department. As can be seen in the middle part of

Figure 4.2, the interviewer dressed more informally than in the Law School condition and more formally than in the Regional Community condition, wearing clothing such as a pair of neat jeans, a plain sweater, and neutral shoes.

Figure 4.2. The interviewer manipulation of Study 4.2 illustrated: The interviewer introduced as coming from Leiden University Law School (left), Leiden University Psychology Department (middle), or Leiden Regional Community College (right).

Procedure. The experimenter went out to train stations (Utrecht and Leiden Central Stations) and at these stations interviewed people from various backgrounds and with different educational levels about their trust in judges. These train stations are major hubs of the Dutch railway system where people from different social demographic backgrounds pass through. The location of data collection did not affect the results reported.

The experimenter visited each station in all three interviewer conditions. The experimental procedure and sampling procedures were the same as in Study 4.1.

Response rates were again similar across interviewer conditions, suggesting that we were successful in minimizing selection effects. Debriefing interviews again suggested that we were successful in our aim to reduce experimenter demands in the study.

Trust in judges. Our main dependent variable was a measure of trust in the judges in the Netherlands. We slightly varied how we operationalized this construct compared to Study 4.1, this time focusing only on the level to which participants feel that judges represent common citizens, grounded on Tyler and Huo (2002). Our measure consisted of the following items: "Dutch judges protect the interests of common people," "Dutch judges understand the problems of common people," and "Dutch judges understand what common people find important." Participants' answers to these items were internally consistent ($\alpha = .88$) and average scores were used for interpretation, with higher scores indicating higher levels of trust in judges.

Results

Trust in judges. To test our predictions, we performed a 3 (interviewer: Law School, Psychology Department, Regional Community College) x 3 (educational level: low, intermediate, high) ANOVA on participants' ratings of trust in judges. This analysis revealed a significant main effect of the interviewer manipulation, $F(2, 127) = 3.13$, $p < .05$, $\eta_p^2 = .05$, a significant main effect of educational level, $F(2, 127) = 8.56$, $p < .001$, $\eta_p^2 = .12$, and a significant interaction effect between the interviewer manipulation and educational level, $F(4, 127) = 2.62$, $p < .04$, $\eta_p^2 = .08$. To interpret these effects we performed a post-hoc Tukey-b test for multiple comparisons between the means ($p = .05$) with the nine cells of our design serving as the independent variable (Kirk, 1982). The subscripts reported in Table 4.2 show the results of this test. Table 4.2 also presents the means and standard deviations of participants' ratings of trust in judges.

As can be seen in Table 4.2, the Tukey-b test showed that when interviewed by the Law School or Psychology Department interviewers, trust in judges among lower educated participants was as high as among the participants with intermediate or higher educational levels. However, comparable to Study 4.1, lower educated participants indicated significantly lower levels of trust in judges to the Regional Community College interviewer than to the Law School interviewer.

Table 4.2 further reveals that the trust ratings that lower educated participants indicated when interviewed by the person from the Psychology Department were closer to those obtained with the Law School interviewer than with the Regional Community College interviewer. We explicitly note, however, that the difference between the Psychology and Regional Community College interviewer conditions

Table 4.2.

Trust in judges as a function of the interviewer being introduced as coming from Leiden University Law School, Leiden University Psychology Department, or Leiden Regional Community College and participants' educational level (Study 4.2)

	Educational Level					
	Low		Intermediate		High	
Interviewer	*M*	*SD*	*M*	*SD*	*M*	*SD*
Law School	$^{x}4.16_{a}$	1.18	4.93_{a}	0.90	4.12_{a}	0.86
Psychology Department	$^{x}3.93_{a,b}$	0.28	4.61_{a}	0.74	4.86_{a}	0.92
Regional Community College	$^{y}2.97_{b}$	0.85	4.43_{a}	1.20	4.38_{a}	1.19

Note. Trust ratings are on 7-point scales with higher values indicating higher levels of trust in Dutch judges. Means with no subscript in common differ significantly ($p < .05$), as indicated by a Tukey-b post-hoc test for multiple comparisons between means among the nine cells of our design. The cells with different superscripts differ significantly ($p < .001$), as shown in direct contrast tests performed against the overall error term in our 3 x 3 design (Kirk, 1982).

among the low-educated participants was not statistically significant in the Tukey-b test. We think this latter aspect of our findings may be caused in part by the fact that although the Tukey-b posthoc test that we report has nice qualities in that it allowed us to examine with one test how the interaction pattern in our 3 x 3 design looked like, the post-hoc quality of the test also had the potential drawback of testing our predictions in a conservative and a-directional manner. We therefore complemented the Tukey-b test with a contrast test that directly tested the difference between the two noted cells (the Psychology and Regional Community cells among the lower educated participants). We tested this contrast against the overall error term in our 3 x 3 design. This contrast test yielded a significant effect, $F(1, 134) = 12.60$, $p < .001$, $\eta_p^2 = .09$, indicating that according to this test lower educated participants responded with significantly less trust toward the interviewer from the Regional Community College than toward the Psychology Department interviewer. Another contrast test showed lower educated participants did not respond differently in terms of trust toward the Law School and Psychology interviewers, $F(1, 134) = 0.14$, $p = .71$, $\eta_p^2 = .00$. The superscripts in Table 4.2 report the results of these contrast tests.

Discussion

In summary, as in Study 4.1, the interviewer manipulation had the strongest effects among lower educated participants. For participants with intermediate and high educational levels, trust in judges was not significantly different depending on how the interviewer presented herself. When interviewed by the Law School interviewer, and

to some extent the Psychology interviewer, lower educated participants indicated levels of trust in judges that were as high as among participants with intermediate or higher education. On the other hand, when interviewed by the non-university interviewer of the Regional Community College the lower educated participants indicated significantly lower levels of trust in judges.

Study 4.3

The aim of Study 4.3 was to build and extend on Studies 4.1 and 4.2 by examining the effects of a different interviewer manipulation among another group of under-investigated participants, namely self-identified female Muslims. Survey studies about trust in law and society are mostly conducted by interviewers from majority groups. In Western countries this means that survey studies about trust in law and society are often conducted by interviewers who do not wear a headscarf. And we again note that if interviewers from minority backgrounds are included in trust studies they tend not be part of the design of the studies. Thus, we varied in Study 4.3 whether the same interviewer presented herself without or with a headscarf to female Muslim participants.

In this study, we examined societal issues for which we thought that female Muslims' responses might naturally be influenced by the majority group identity of interviewers. More specifically, in Study 4.3 we focus on female Muslims' responses to questions about trust in society, trust in non-Muslims (which is the majority group in the Netherlands), and evaluations of non-Muslims. In this way we examined the question whether female Muslims indicate different levels of trust in Dutch society and trust in non-Muslims toward the interviewer without a headscarf compared to when the same interviewer presents herself with a headscarf.

Method

Participants and design. One hundred-twenty Muslim women participated voluntarily in the study. Participants were interviewed by the same interviewer presenting herself without a headscarf (N = 60) or with a headscarf (N = 60). Of the participants, 100 were born in the Netherlands, 5 in Morocco, 3 in Turkey and 11 in other countries (1 score was missing). Of the fathers of the participants, 4 were born in the Netherlands, 70 in Morocco, 27 in Turkey and 17 somewhere else (2 scores were missing). Of the mothers of the participants, 7 were born in the Netherlands, 69 in Morocco, 25 in Turkey, 17 somewhere else (2 scores were missing). The average age was 21.28 year (SD = 3.56, range: 16-33 years). Data were collected on weekdays between February 3 and February 16, 2015.

The key prediction in this study was an effect of the interviewer manipulation among a group of under-investigated participants (in this case female Muslims). The sample size among this group of minority participants was grounded by an a-priori G power

analysis (Faul et al., 2007) with $\alpha = 0.05$, power = .95, $\eta_p^2 = .10$, based on the simple main effect of the interviewer manipulation within the group of under-investigated participants in Study 4.2 (in that study lower educated participants).

Experimental procedure.

Interviewer manipulation. We manipulated whether the same interviewer either wore a headscarf or did not wear a headscarf, as can be seen in Figure 4.3.

Procedure. Our interviewer went out to shopping malls and interviewed self-identified female Muslims. Data were collected in shopping malls in Utrecht, Eindhoven, and Rotterdam. People from different social demographic backgrounds visited the malls. The location of data collection did not affect the results reported.

The interviewer visited every mall both in the headscarf and no-headscarf condition, to keep sampling procedures between the conditions similar. Female Muslims were

Figure 4.3. The interviewer manipulation of Study 4.3 illustrated: The interviewer of Study 4.3 wearing no headscarf (left) or a headscarf (right).

initially identified by the interviewer on the basis of their headscarf or Moroccan or Turkish facial features, as Moroccans and Turks constitute the biggest Muslim groups in the Netherlands. Each person thus identified was asked to participate without distinguishing on apparent educational level or age. Response rates did not differ between interviewer conditions.

The experimental procedure was the same as in Studies 4.1 and 4.2 to rule out possible experimenter demands. For example, the experimenter used standard pre-scripted communication in the interaction with participants. And, as in Studies 4.1 and 4.2, participants were given envelopes with the structured, paper-and-pencil questionnaires inside, and while the interviewer kept a distance, participants completed the questionnaires on their own, put the questionnaires in the envelopes, sealed the envelopes, and only then returned the envelopes to the interviewer.

Trust in Dutch society. Five questions measuring general trust in people were based on widely used societal trust scales (Reeskens & Hooghe, 2008; Rosenberg, 1956, 1957; Zmerli & Newton, 2008): "I have the feeling that most Dutch people can be trusted," "I have the feeling that most Dutch people would try to exploit you if they got a chance," "I have the feeling that most Dutch people try to help other people if they got a chance," "I have the feeling that most Dutch people are in general good people," and "I have trust in Dutch society." The reliability of the resulting 5-item scale was deemed sufficient for theory-testing purposes (α = .67; see Nunnally & Bernstein, 1994). Average scores were used for interpretation, and higher scores indicated more trust. Unless otherwise specified, all ratings in Study 4.3 were measured on 7-point-Likert scales (1 = *completely disagree,* 7 = *completely agree*).

Trust in Non-Muslims. Based on our analysis of literature on social trust (e.g., Cehajic, Brown & Castano, 2008; Gundelach, 2014; Rotter, 1967; Tam, Hewstone, Kenworthy, & Cairns, 2009) and the specific context we are studying here, we assessed trust in non-Muslims by asking participants to what extent they agreed with the following statements: "Despite earlier racism against Muslims, I do trust non-Muslims," "I trust non-Muslims when they say they want to live in peace with us," "I do not trust them because they want revenge for the actions of Muslim-extremists," "Until non-Muslims have proven that they can be trusted, I think that you should pay attention when you interact with them," "I have positive expectations when I have contact with non-Muslims," "I have the feeling that non-Muslims would misbehave against Muslims if they would not be punished for this," and "I trust Muslims more than non-Muslims." The items (N = 7) were sufficiently associated with each other (α = .70). Average scores were used for interpretation, with higher scores indicating higher levels of trust in non-Muslims.

Positive evaluations of Non-Muslims. The Individualized Trust Scale (Wheeless & Grotz, 1977) is another way of measuring trust towards groups. It measures to what extent groups rated possess 15 characteristics (e.g., exploiting, straightforward), demonstrated to be related to the perceived trustworthiness (Berlo, Lemert & Mertz, 1969) and credibility of these groups (McCroskey, Jensen, & Valencia, 1973; McCroskey, Scott, & Young, 1971). The 8 positive and 7 negative characteristics in the scale were alternated in how we presented them to our participants. Average scores were used for interpretation, with higher scores indicating more positive evaluations of non-Muslims. The questions were sufficiently associated with each other ($\alpha = .89$).

Results

The main aim of Study 4.3 was to investigate whether female Muslims would respond differently on scales about trust in society toward an interviewer who presented herself with a headscarf compared to when the same interviewer who presented herself without a headscarf. To examine this, our main analyses in Study 4.3 were ANOVAs with the headscarf manipulation as the independent variable and the scores on the scales of trust in Dutch society, trust in non-Muslims, and evaluations of non-Muslims and identification with Dutch society as dependent variables.

Trust in Dutch society. Participants scored significantly higher on the scale measuring trust in Dutch society when the interviewer presented herself without a headscarf ($M = 4.68$, $SD = 0.84$) than when the same interviewer presented herself with a headscarf ($M = 4.25$, $SD = 0.86$), $F(1, 112) = 7.36$, $p = .008$, $\eta_p^2 = .06$. Thus, female Muslim participants expressed significantly more trust in Dutch society in the no-headscarf condition than in the headscarf condition.

Trust in Non-Muslims. Participants scored higher on the scale that assessed trust in non-Muslims when the interviewer presented herself without a headscarf ($M = 4.77$, $SD = 0.88$) than when the same interviewer presented herself with a headscarf ($M = 4.21$, $SD = 0.84$), $F(1, 108) = 11.52$, $p = .001$, $\eta_p^2 = .10$. Hence, participants expressed significantly more trust in non-Muslims in the no-headscarf condition than in the headscarf condition.

Positive evaluations of Non-Muslims. Similarly, participants gave more positive evaluations of non-Muslims when the interviewer presented herself without a headscarf ($M = 4.74$, $SD = 0.93$) than when the same interviewer presented herself with a headscarf ($M = 4.28$, $SD = 0.69$), $F(1, 102) = 6.31$, $p = .014$, $\eta_p^2 = .06$. In other words, participants responded more positively about non-Muslims in the no-headscarf condition than in the headscarf condition.

Thus, in accordance with our predictions, female Muslim participants indicated more trust in Dutch society, more trust in non-Muslims, and more positive evaluations of non-Muslims to an interviewer without a headscarf compared to the same interviewer with a headscarf.

General Discussion

In three field experiments we experimentally investigated the role that interviewers can play in studies on trust in law and society, especially when examining trust in law and society among under-investigated participants. In our first two studies we found that when completing questionnaires on trust in a country's judges as given to them by interviewers who presented themselves as from university law schools, lower educated participants reported that they hold high levels of trust in judges. Specifically, when interviewed by Law School interviewers, trust levels among lower educated participants were as high as among intermediate and highly educated participants. Thus, in these conditions overall trust in the national judges was reported to be high. This corresponds with trust surveys that often report that trust in the national justice system is rather high (e.g., TNS Political & Social, 2013). However, when the same interviewer presented herself as non-university interviewer, trust levels among lower educated participants were significantly lower. In these conditions, lower educated participants indicated much less trust in judges. When the lower-educated participants were interviewed by an interviewer from a Psychology Department the trust ratings were somewhere in between, albeit somewhat closer to the Law School interviewer condition than to the condition in which the interviewer was from the Regional Community College. This fits with our line of reasoning such that it can be assumed that people working at Law Schools are more part of the status quo than those working at Psychology departments. More research is needed to examine the implications of this assumption in detail.

The effects found in our third study among female Muslims followed a similar pattern as the key findings of Studies 4.1 and 4.2. When answering questionnaires on degree of trust in Dutch society and non-Muslims as given to them by an interviewer presenting herself without a headscarf, female Muslim participants indicated higher levels of trust in Dutch society and non-Muslims. Yet, when the same interviewer presented herself with a headscarf, female Muslims participants' responses were significantly different. In fact, female Muslim participants reported significantly less trust in Dutch society and non-Muslims when the interviewer was wearing a headscarf as opposed to not wearing a headscarf.

Hence, all three experimental studies showed that degree of trust in law and society varied depending on the interviewer who handed out the trust questionnaire. That is, in line with our predictions, our systematically controlled differences between interviewers

had reliable effects on the degree of trust in law and society among female Muslim participants (Study 4.3) and among lower educated participants (Studies 4.1 and 4.2). Participants with intermediate and high educational levels, on the other hand, were not significantly affected by our interviewer manipulations (Studies 4.1 and 4.2). Thus, interestingly, most of the action of our interviewer manipulations took place among groups of participants that are currently under-investigated in most behavioral sciences studies (Henrich at al., 2010).

An interviewer with a non-university Regional Community College affiliation has a lower educated and less rich background than an interviewer with a Law School affiliation. An interviewer wearing a headscarf may be seen as having a less Western (ethnic or religious) background than an interviewer not wearing a headscarf. We emphasize that female Muslims in Western societies are not necessarily from non-Western or non-WEIRD backgrounds. This noted, we also think it is reasonable to conclude that systematically incorporating the concept of "WEIRD-ness" in experiments and other empirical studies can reveal important insights into the issues pertaining to trust and law and can lead to the observation of quite different levels of trust, particularly when interviewing lower educated participants.

Limitations and Future Research

The fact that especially the lower educated participants apparently were tempted to go along with what they were thinking the university affiliated interviewer was conceiving of the issue of trust in judges (after all, highly educated people tend to have higher levels of trust in law; e.g., TNS Political & Social, 2013) is interesting and, in fact, precisely tests our line of reasoning that WEIRD interviewers affect trust reporting among under-investigated participants. Our observation that only in case of the non-university interviewer the lower educated participants gave back the questionnaires to the interviewer with statements such as "You know, these judges are really different from you and me" underscores our line of thought about the important role interviewers can play when examining trust in law and society among understudied groups.

Similarly, the fact that female Muslim participants went along with higher ratings of trust in Dutch society and non-Muslims in case the interviewer was not (as opposed to was) wearing a headscarf fits with our line of reasoning. After all, non-Muslims in the Netherlands tend to indicate higher trust in the country than Muslims in the Netherlands tend to do (Huijnk et al., 2015), and trust levels of female Muslims were also quite high in case of the headscarf-absent interviewer (at least significantly higher than in case of the headscarf-present interviewer). Also, only in case of the headscarf-wearing interviewer did we find that participants gave back the envelopes with statements such as "How the Dutch Police is treating Moroccan guys is just not right",

which also adheres to our assumption that interviewer differences play a role when examining trust in law and society among members of minority groups. This noted, of course future research needs to be done to follow up the results presented here.

Importantly, please recall that we ensured in all our studies that participants completed the questionnaires on their own without the experimenter being in their immediate presence. Furthermore, participants received the questionnaires in envelopes and after filling out the questionnaires put these back into the envelopes, sealed the envelopes, and only then gave the materials back to the experimenter. In short, we tried our utmost best to reduce experimenter demands (Orne, 1962) in the current studies, and debriefing suggested that we were successful in doing this.

In Studies 4.1 and 4.2 we crossed the interviewer manipulations with the three qualitatively different educational levels in the Netherlands (CBS, 2008). It may be interesting to examine how interviewer manipulations like the ones we used in the first two studies would affect trust reporting when more than three educational categories can be distinguished in future research studies outside the Netherlands. The inclusion of male interviewers would also be an important extension of the studies presented here.

We took ample measures to minimize that selection processes were present in our studies, and found no indications that such processes were affecting the differences in trust reporting that we found between interviewer conditions. For example, if female Muslim participants would not have wanted to participate in the research if a non-Muslim interviewer would have asked them, it would seem more likely that these participants would distrust and dislike non-Muslims as well. (In fact, participants in the headscarf condition in Study 4.3 still scored rather high on trust in non-Muslims and positive evaluations of non-Muslims.) Nevertheless, selection processes can never be fully excluded in field experiments like the studies reported here. If selection processes would be shown to be robust in future research projects, then this perhaps could also be treated as evidence for the interesting interviewer effects in the domains of trust in law and society that we think our studies have brought to light. That is, the practical implications of our studies would still be the same: Studies on trust in law and society may be missing crucial patterns in degree of trust among under-investigated participants. Thus, if selection processes would play a role the fact that these studies are often conducted by WEIRD interviewers might lead to the under-representation of already under-investigated participants. To further examine the influence of selection processes future research studies might keep the presentation of the interviewer similar and vary only in the paper-and-pencil questionnaire whether the questionnaire is coming from a university or non-university type of organization. Thus, for example, future research could explore whether lower educated participants report different

levels of trust in judges merely depending on where the trust questionnaire is coming from (such as Law Schools versus Community Colleges).

In real-world interpersonal exchanges between interviewers and participants the impression that the interviewer belongs to a different social category or group may be triggered by a mixture of many different things, such as the interviewer's affiliation, style of clothing, and way of talking (e.g., accent, choice of words). In Studies 4.1 and 4.2, we were interested whether lower educated participants would respond differently when the interviewer comes from a lower educational background as opposed to higher educational background. Thus, corresponding to the situation of participating in a survey in real-life, in Studies 4.1 and 4.2 we controlled the interviewer's affiliation, name, and dress style to reflect educational background. We suggest that this may add to the ecological validity of the research. In our studies we approached people in their everyday life, and the information we obtained is influenced by real-life concerns. Moreover, we followed the recommendation by John Thibaut, a pioneer researcher and expert social scientist in the domain of law and society: *"If you investigate something new, start with a strong manipulation of your independent variable!"* (E. A. Lind, personal communication, May 6, 2015). In Study 4.3, we only varied whether the interviewer was or was not wearing a headscarf. While a lot of field studies lack experimental control, the experimental manipulations in our studies have face validity we think and give us the opportunity to determine causal effects of the interviewer on trust reporting among both lower educated people and female Muslims.

In the current studies, we used carefully trained interviewers to make sure that interviewees experienced them as indeed coming from the university or non-university organizations or majority or minority groups. In doing so, we also learned from debriefing interviews following one pretest study that the individual experimenter in that pilot study was not very successful in achieving that the university and non-university interviewer conditions were psychologically real to participants. The current set of three studies presented here was successful in achieving these aims, as also our debriefing interviews indicated, and we note that careful training of interviewers and thorough pilot testing is needed before conducting these kinds of non-laboratory experiments.

Implications and Conclusions

In the current field experiments we found that lower educated people and female Muslims indicated different degrees of trust in an important legal institution like the judiciary and different degrees of trust in the broader society as a whole depending on who interviewed them. These findings may have important implications for how research studies on trust in law and society are often conducted.

Interviewers who carry out studies on trust in law and society are quite often higher educated people and members of the majority group. Thus, interviewers who examine trust in law and society will often come from WEIRD backgrounds. Furthermore, as noted, when interviewers do come from non-WEIRD backgrounds, then this aspect is not systematically addressed in the research design of the trust studies involved.

The current findings suggest that practices of how interviewers on trust in law and society commonly present themselves can have an effect on levels of trust in law and society, especially among understudied participants. For example, interviewers may tend to introduce themselves properly by mentioning the organization responsible for the research. For instance, we as university researchers may often introduce ourselves with our university affiliation. In addition, interviewers may tend to dress up properly when interviewing people, at least interviewers from for instance a Law School will often do so. Furthermore, most interviewers will belong to the majority group and in Western countries this means that most interviewers will thus not be wearing a headscarf.

Thus, in the interviewer conditions that represented how trust research is quite often conducted (i.e., university interviewers and interviewers without headscarves) we found relatively high levels of trust in law and society. In fact, in these interviewer conditions, trust levels among the lower educated were as high as among the middle and highly educated. However, in the conditions that represented how trust studies are less often conducted (i.e., non-university and headscarf interviewers) trust scores among lower educated people and female Muslims were significantly lower.

In other words, an important implication of the current findings is that lower levels of trust may well be missed because of the way trust studies are often conducted, especially when examining trust among under-investigated participants. In many large-scale trust surveys, trust in for instance the national judges tends to be relatively high, at least in many European countries (e.g., TNS Political & Social, 2013). If lower levels of trust in law and society would indeed be missed in many studies of public trust, society's so-called social capital may not be assessed in an accurate manner (De Tocqueville, 1945). After all, reports of public trust in law and society are used by governments and societal institutions, and trust reporting might translate into policy. This further illustrates the importance of obtaining accurate measures of levels of trust in law and society. When crucial patterns of trust in mainstream institutions and the broader society as a whole would indeed be missed, as our findings suggest, conclusions about the state of democracy of society and about the legitimacy of law and society may be flawed or at least not as accurate as would be preferable.

We explicitly note here that the current studies were not meant to examine whether levels of trust among some participants are concealed or hidden from WEIRD interviewers. That is not what this research is about, we think. What we do learn from

the present research is that levels of trust in law and society among some participants vary quite substantially depending on the interviewer conducting the trust study.

The current research has important implications for the methodology of studies on trust in law and society and the conceptual conclusions that are based on these studies. To us, the current findings suggest that there is some value to more systematically and rigorously re-evaluate the designs of our studies on trust in law and society. Our findings indicate that experimental research methods can contribute to examining the key role of interviewers in trust studies. The interviewer effects that we found in our studies indicate that the background of interviewers should be systematically addressed in the research design of studies on trust in law and society. Thus our research suggests that even for trust researchers who are more interested in using survey or interview methods, the background of interviewers should be systematically included in their research designs. For example, when selecting interviewers to conduct a survey, researchers might want to consider whether certain issues of group membership or societal categorization may play a role in the eyes of at least some research respondents. For instance, studies about radicalization might make Muslim participants become aware of their religious group membership, which in turn could motivate them to respond in a certain way depending on who is interviewing them.

It seems reasonable to assume that there is less social distance between a non-university interviewer and lower educated participants than between a university interviewer and lower educated participants in terms of educational level. Similarly, there is less social distance between an interviewer with a headscarf and female Muslim participants than between an interviewer with a headscarf and female Muslim participants in terms of religious or ethnic background. Viewed in this way, the implication of our findings is that social distance between interviewer and participant has an effect on levels of trust in law and society, at least among some participants. That is, when our interviewer manipulations produced heightened social distance between interviewer and participant, lower educated participants and female Muslim participants reported high levels of trust in law and society. However, when our interviewer manipulations created lower social distance between the interviewer and these participants, these participants reported significantly lower levels of trust in law and society. We would applaud future research that focuses on this possible implication of our findings.

Future research could also investigate the issue of interviewer effects in studies about figures or social movements that are not part of the societal status quo, such as support for extreme right-wing political parties, anti-globalization movements, or terrorist organizations. If the issue of interviewer effects would turn out to be important in other research domains, then this could also imply that opinion researchers may easily miss how under-represented citizens will vote for political candidates or support social movements that are not part of the societal status quo.

Conclusions

Speculation aside, the current field experiments have provided support for our line of reasoning that the organizational affiliation and group identity of interviewers affect trust reporting among under-investigated participants, such as those with lower education or from Muslim backgrounds. Thus, we can conclude that the issue of interviewer effects on trust reporting among under-investigated participants is indeed a crucial topic when studying trust in law and society. The current research offers a promising proposition for how trust in law and society can be studied in controlled ways. In short, our findings suggest that it is important to experimentally investigate the critical role that interviewers play in studies of trust in law and society, especially when examining under-investigated participants.